John von Neumann Institute for Computing

Stretching RNA Hairpins

C. Hyeon, D. Thirumalai

published in

From Computational Biophysics to Systems Biology (CBSB07),
Proceedings of the NIC Workshop 2007,
Ulrich H. E. Hansmann, Jan Meinke, Sandipan Mohanty,
Olav Zimmermann (Editors),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **36**, ISBN 978-3-9810843-2-0, pp. 61-68, 2007.

http://www.fz-juelich.de/nic-series/volume36

Stretching RNA Hairpins

Changbong Hyeon[1] **and D. Thirumalai**[2,3]

[1] Center for Theoretical Biological Physics, University of California at San Diego
La Jolla, California 92093, USA

[2] Biophysics Program
Institute for Physical Science and Technology

[3] Department of Chemistry and Biochemistry, University of Maryland
College Park, MD 20742, USA
E-mail: thirum@glue.umd.edu

1 Introduction

Optical tweezer experiments have used mechanical force to trigger folding and unfolding of RNA molecules at a single molecule level[1,2]. Application of tension to a specific position of the molecule, induces sequence and structure-dependent response. The measured force-extension curve (FEC) is fit using the worm-like (WLC) model[3,4]. The stability of RNAs is inferred by integrating the FECs. For simple motifs, such as hairpins, it has been shown that the stability of the native structures can be accurately measured using mechanical unfolding trajectories which exhibit multiple transitions between the folded and the unfolded state when the force is held constant[1].

Mechanical force has also been used to probe unfolding and refolding kinetics of RNA. The cooperative reversible folding of hairpins has been shown by monitoring the end-to-end distance (R), a variable conjugate to the mechanical f, as a function of time. This procedure works best when RNA folding is described using two-state approximation. For multidomain ribozymes the folding/unfolding kinetics is complex and new tools are required to interpret the kinetic data. In a pioneering study Onoa *et. al.*[2] showed that the rips in FECs for the L-21 derivative of *Tetrahymena thermophila* ribozyme (*T.* ribozyme), composed of multiple domains, are a result of unfolding of individual intact domains that are stabilized in the native state by counterion-dependent tertiary interactions.

In this note, we probe the forced-unfolding dynamics of RNA hairpins using a simple model. Because these simulations can be used to directly monitor structures in the transition from folded to the fully stretched states, unfolding pathways can be unambiguously resolved. We use a *Self-Organized Polymer (SOP) model* for RNA that is based only on the self-avoiding nature of the RNA and the native structure. We apply the SOP model to probe forced-unfolding of a number of RNA structures of varying complexity. Many of the subtle features of the variations in the mechanical unfolding as a function of f and r_f can be illustrated using P5GA, a simple RNA. Our results show that the response of RNA to force is largely dependent on the architecture of the native state. More importantly, we have established that the characterization of the the energy landscape requires using force values (or loading rates) over a wide range.

2 Methods

Model: The SOP model for RNA that retains chain connectivity and favorable attractive interactions between sites that stabilize the native fold. Each interaction center represents the center of mass of a nucleotide. In terms of the coordinates $\{\mathbf{r}_i, i = 1, 2, \ldots N\}$ of RNA with N nucleotide the total potential energy in the SOP representation is

$$\begin{aligned} V_T &= V_{FENE} + V_{nb}^{(att)} + V_{nb}^{(rep)} \\ &= -\sum_{i=1}^{N-1} \frac{k}{2} R_0^2 \log(1 - \frac{(r_{i,i+1} - r_{i,i+1}^o)^2}{R_0^2}) \\ &+ \sum_{i=1}^{N-3} \sum_{j=i+3}^{N} \epsilon_h [(\frac{r_{ij}^o}{r_{ij}})^{12} - 2(\frac{r_{ij}^o}{r_{ij}})^6] \Delta_{ij} \\ &+ \sum_{i=1}^{N-2} \epsilon_l (\frac{\sigma^*}{r_{i,i+2}})^6 + \sum_{i=1}^{N-3} \sum_{j=i+3}^{N} \epsilon_l (\frac{\sigma}{r_{ij}})^6 (1 - \Delta_{ij}). \end{aligned} \quad (1)$$

The first term is for the chain connectivity. The finite extensible nonlinear elastic (FENE) potential[5] is used with $k = 20 kcal/(mol{\cdot}\text{Å}^2)$, $R_0 = 0.2\ nm$, and $r_{i,i+1}$ is the distance between neighboring beads interaction centers i and $i + 1$, $r_{i,i+1}^o$ is the distance in the native structure. The use of FENE potential is more advantageous than the standard harmonic potential especially when considering forced-stretching because the fluctuations of $r_{i,i+1}$ are strictly restricted around $r_{i,i+1}^o$ with variation of $\pm R_0$. The Lennard-Jones potential is used to account for interactions that stabilize the native topology. Native contact is defined for the pair of interaction centers whose distance is less than $R_C = 1.4$ nm in the native state for $|i - j| > 2$. If i and j sites are in contact in the native state, $\Delta_{ij} = 1$, otherwise $\Delta_{ij} = 0$. We used $\epsilon_h = 0.7\ kcal/mol$ for the native pairs, $\epsilon_l = 1$ $kcal/mol$ for non-native pairs. In the current version, we have neglected non-native interactions which will not qualitatively affect the results because, under tension, such interactions are greatly destabilized. To ensure the non-crossing of the chain, we set $\sigma = 7$Å. Only for $i, i + 2$ pairs we set $\sigma^* = 3.5$ Å to prevent the flattening of the helical structures when the overall repulsion is large. There are five parameters in the SOP force field (k, R_0, ϵ_h, ϵ_l, and R_c)[6]. Of these the results are sensitive to the precise values of ϵ_h/ϵ_l and R_c. We have discovered that the quantitative results are insensitive to R_c as long as it is in the physical range that is determined by the RNA contact maps. In principle, the ratio ϵ_h/ϵ_l can be adjusted to obtain realistic values of forces. For simplicity we choose a uniform value of ϵ_h for all RNA constructs. Surprisingly, the SOP force field, with the same set of parameters, can be used to obtain near quantitative results for RNA molecules of varying native topology.

Simulations: Using the SOP model, we simulated the mechanical unfolding and refolding of P5GA hairpin. In force-clamp simulations a constant force is applied to one end of the molecule while the other end is fixed. Finally, in force-quench computations the force on the molecule is reduced to the final value to initiate mechanical refolding. In both force-clamp and force-quench setup the dynamics of the linker (usually hybrid RNA/DNA handles) is not relevant whereas depending on the characteristics of the linkers the dynamics of linker may play an important role in the force-ramp experiments[7].

Time scales: Since a typical value for the mass of a nucleotide, $m \sim 300g/mol - 400g/mol$, the average distance between the adjacent nucleotides in the SOP representation of RNA is $a \approx 5$ Å, $\epsilon_h = 0.7\ kcal/mol$, the natural time is $\tau_L = (\frac{ma^2}{\epsilon_h})^{1/2} = 3 \sim 5ps$. We use $\tau_L = 4.0ps$ to convert simulation times into real times. To estimate the time scale for mechanical unfolding dynamics we use a Brownian dynamics algorithm[8,9] for which the natural time for the overdamped motion is $\tau_H = \frac{\zeta\epsilon_h}{k_B T} h\tau_L$ ($\tau_L = 4ps$). We used $\zeta = 100\tau_L^{-1}$ in the overdamped limit, that approximately corresponds to friction constant of a nucleotide in water. The equations of motion in the overdamped limit are integrated using the Brownian dynamics algorithm.

Force-induced transitions in a simple hairpin (P5GA): Liphart *et. al.* showed that the P5ab hairpin, the construct in which P5c stem-loop and the A-rich bulge in P5a are removed from the P5abc subdomain in T. ribozyme, reversibly folds in an all-or-none fashion upon application of constant force. The equilibrium between the native basin of attraction (**NBA**) and the unfolded basin of attraction (**UBA**) can be shifted by altering the value of the constant force, f_c. To probe the two-state behavior of hairpins under force we used a smaller 22-nt hairpin, P5GA (Protein Data Bank (PDB) id: 1eor)[10,11]. For the P5GA hairpin simulations over a wide range of forces can be performed in reasonable times. The topologically simple hairpin has a single tetra-loop and nine consecutive base pairs. In an earlier study[11] we showed, using a minimal three interaction site (TIS) model in which each nucleotide is represented by three sites, that the dynamical behavior of P5GA under tension is qualitatively similar to P5ab. The much simpler SOP representation of P5GA allows us to probe exhaustively the folding and unfolding kinetics of the hairpin that is manipulated by force-ramp, force-quench, and force-clamp.

The hallmark of P5ab[1] and P5GA[11], when a constant force is applied to either the 3' or the 5' ends, is the observation of bistable kinetics. When a constant f_c is applied to the 3' end P5GA makes transitions (Fig. 1-B) between the **UBA** ($R \approx 8nm$) to the **NBA** ($R \approx 2nm$). At $f_c =$ 14.0, 15.4, 17.5 pN a large number of transitions occur over $45ms$ duration which suggests that the hairpin dynamics is effectively ergodic. As in our previous study[11] the equilibrium constant between the folded and unfolded hairpin calculated using a long mechanical unfolding trajectory coincides with an independent ensemble average calculation i.e., time averages are roughly equivalent to ensemble averages. When $f_c = 14\,pN$ the residence time in the **NBA** is much greater than in the **UBA** while at $f_c = 16.8\ pN$ the **UBA** is preferentially populated (Fig. 1-B). The population of P5GA in the **NBA** changes when f_c is varied can be seen in the histogram ($P(R)$) of the end-to-end distance R (Fig. 1). At $f_c = 15.4\,pN$, which is slightly above the midpoint of the **NBA** $\Leftrightarrow$ **UBA** transition, several jumps between the **NBA** and **UBA** are observed. The $P(R)$ distribution reflects the bistable nature of the landscape. The free energy profile with respect to R is computed using $\Delta F(R) = -k_B T \log P(R)$. From $P(R)$ at $f_c = 15.4$ we can obtain the free energy of stability of the folded hairpin with respect to the unfolded state using $\Delta G \approx f_c \Delta R_{UF}$ where ΔR_{UF} is the distance between the folded and unfolded states of P5GA. Using $\Delta R_{UF} \approx 6nm$ we find that $\Delta G \approx 13kcal/mol$. The Vienna RNA package[12], which uses an entirely different free energy parameters for RNA, gives $\Delta G \approx 12.8kcal/mol$. This comparison shows that the SOP model can, for simple structures, give accurate results for stability. At $f_c = 15.4\,pN$, the transition barrier is about $\sim 1.5k_B T$. The **UBA** is more populated at this value of f_c.

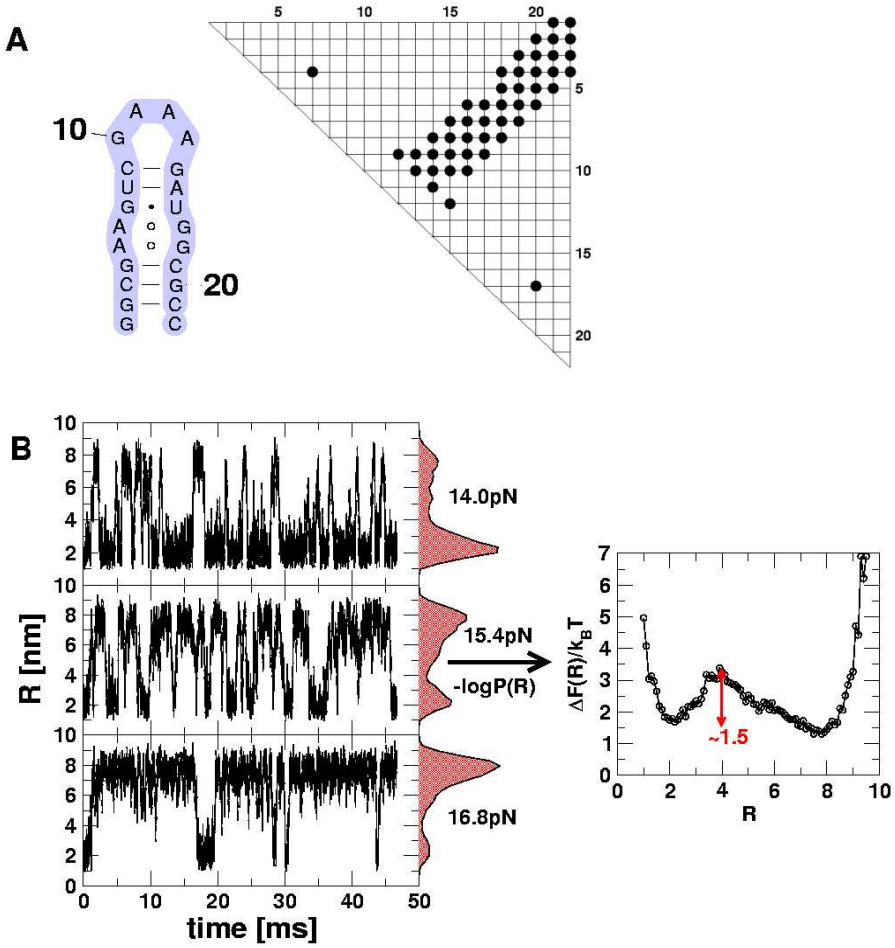

Figure 1. **A.** The secondary structure of P5GA hairpin and its contact map. **B.** The time-dependent fluctuations of P5GA hairpin between the folded ($R \approx 1.5nm$) and unfolded ($R \approx 8nm$) states. The end-to-end distance changes spontaneously between two values. The force-clamped dynamics of the P5GA hairpin is probed for $\sim 45\ ms$. The histograms $P(R)$ at $f_c = 14.0, 15.4, 16.8\ pN$ are shown in red. The free energy profile $\Delta F(R)$ as a function of R for $f_c = 15.4\ pN$ on the right shows two-state behavior.

The observed transition times are much shorter than the residence times in each basin of attraction which is also a reflection of the underlying cooperativity of the all-or-none of nature of hopping between **UBA** and **NBA**.

Force-ramp: We also performed force-ramp simulations by subjecting the P5GA hairpin to a continuously changing force, i.e., varying the loading rate (Methods). The simplicity of the SOP model allows us to use values of r_f that are comparable to those used in LOT experiments. At $r_f = 45\ pN/s$ ($\sim 10 r_f^{LOT}$) the force-extension curves show a transition

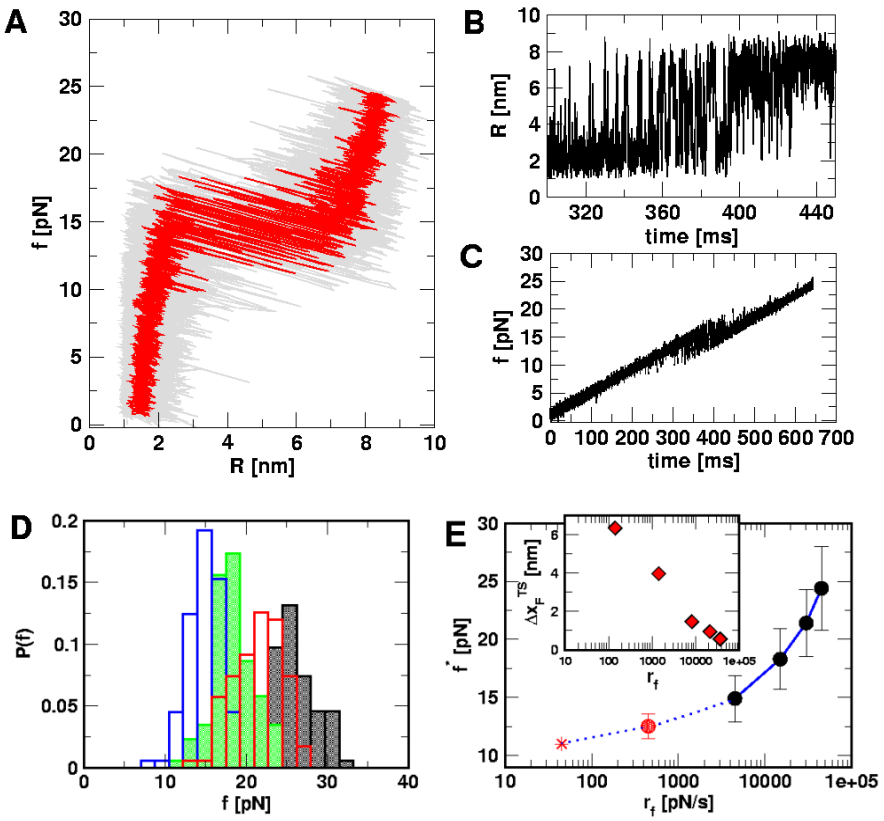

Figure 2. Force-ramp unfolding of P5GA hairpin. **A.** An example of FEC at the loading rate $r_f = 45\ pN/s$ ($k_s = 0.07\ pN/nm$, $v = 0.64\ \mu m/s$). The data points are recorded every 50 μs (grey color), but for better illustration the running average is displayed every 500 μs (red color). **B.** The end-to-end distance (R) as a function of time. **C.** Time dependence of f. The loading rate $r_f = df/dt$ is nearly a constant. **D.** Distribution of unbinding forces from 100 trajectories at four loading rates $r_f = r_f^o$, $\frac{10}{3}r_f^o$, $\frac{20}{3}r_f^o$, and $10r_f^o$ ($r_f^o = 4.5 \times 10^3$ pN/s ($k_s = 0.7\ pN/nm$, $v = 6.4\ \mu m/s$)). The red star is the rupture force value ($f^* = 12.5\ pN$) at $r_f = 45$ pN/s. **E.** Plot f^*, the most probable unfolding force, as a function of r_f. The position of transition state Δx_F^{TS} is computed using $\Delta x_F^{TS} = k_B T \frac{\Delta \log r_f}{\Delta f^*}$. The inset shows the variation of Δx_F^{TS} as a function of r_f.

to the **UBA** at around $f \sim 13\ pN$ (Fig. 2-A). As the force dynamically increases we observe bistable fluctuations in the FEC between the **NBA** and the **UBA** just as when force is held constant (Fig. 2-A). The conformational fluctuations between the two states are unambiguously seen in the time-dependence of the end-to-end distance ($R(t)$) (Fig. 2-B). As time progresses the force is ramped up which results in global unfolding ($R \approx 8\ nm$) for $t > 400\ ms$ (Fig. 2-B). During the time scale of simulation we find frequent and sharp transition between the **UBA** and the **NBA** (Fig. 2-B).

The location of the unfolding transition state Δx_F^{TS} for proteins and RNA is often estimated from force-ramp experiments using the variation of the most probable rupture

force with r_f ($[f^*, \log r_f]$ plot). The loading rate, $r_f = df(t)/dt$, and can be accurately estimated from the slope of the time dependence of $f(t)$ as a function of time. The slope of $f(t)$ as a function of t (Fig. 2-C) is nearly the same as $r_f \approx k_s \times v$, where v is the pulling speed. Strictly speaking, $r_f = k_{eff} \times v$ with $k_{eff}^{-1} = k_s^{-1} + k_{mol}^{-1} + k_{linker}^{-1}$ and k_s, k_{mol}, and k_{linker} are the spring constants of the optical trap, the RNA molecule, and linker, respectively. Typically $k_s \ll k_{mol}, k_{linker}$, thus $k_{eff} \approx k_s$[13]. Throughout the paper we obtain the loading rate using $r_f = k_s v$.

From the force distributions, computed at four different loading rates (Fig. 2-D), we observe that the most probable rupture force (f^*) does not increase logarithmically over a wide range of loading rates (Fig. 2-E). Only if the range of r_f is restricted f^* changes linearly with r_f[14]. The location of the transition state (Δx_F^{TS}) is usually calculated using $f^* \approx (\frac{k_B T}{\Delta x_F^{TS}}) \log r_f$[14] which may be reasonable as long as r_f range is small. However, the $[f^*, \log r_f]$ plot is highly nonlinear (Fig. 2-E). If we use linear regression to analyze the $[f^*, \log r_f]$ plot then $\Delta x_F^{TS} \sim 0.8\ nm$ for the distance between the **NBA** and the transition state. The small value of Δx_F^{TS} is a consequence of the large variation of Δx_F^{TS} as r_f is changed[11,15]. If the loading rate is varied over a broad range, the r_f-dependence of Δx_F^{TS} is manifested as a pronounced convex curvature[7] in the $[f^*, \log r_f]$ plot (Fig. 2-E). Based on the equilibrium free energy profile $F(R)$ as a function of R[11] we expect that for P5GA $\Delta x_F^{TS} \approx 3\ nm$ if r_f is small. Indeed, from the constant force simulation results in Fig. 1-B we find $\Delta x_F^{TS} \approx \frac{R_U - R_F}{2}$. Thus, the slope of $[f^*, \log r_f]$ should decrease (Δx_F^{TS} increases) as r_f decreases. To illustrate the dramatic movement in the transition state we have calculate Δx_F^{TS} using f^* values for two consecutive values of r_f. For example, using $f^* \approx 12.5\ pN$ at $r_f = 450\ pN/s$ and $f^* \approx 14.9\ pN$ at $r_f = 4.5 \times 10^3\ pN/s$ (a value that can be realized in AFM experiments) we obtain $\Delta x_F^{TS} \approx 4.0\ nm$. From the values of f^* at five values of r_f we find that Δx_F^{TS} can move dramatically (Fig. 2-E). In particular, we find that Δx_F^{TS} changes by nearly a factor of ten as the loading rate is decreased to values that are accessible in LOT experiments (see inset in Fig. 2-E). Because of the nearly logarithmic variation of Δx_F^{TS} on r_f over a narrow range of r_f we do not expect Δx_F^{TS} to change appreciably if r_f is lowered from $45pN/s$ to $\approx 5pN/s$. At high r_f or f_c the unfolded state is greatly stabilized compared to the folded state. From Hammond's postulate[16], generalized to mechanical unfolding[7], it follows that as f_c increases Δx_F^{TS} should move closer to the native state. The simulations are, therefore, in accord with the Hammond's postulate.

3 Conclusions

We have used the self-organized polymer representation of RNA hairpins to predict their mechanical unfolding trajectories. Constant force and force-ramp simulations show that dramatic changes in the force profiles take place as the loading rates and the values of the force are varied. If the force is varied over a wide range then regions of the energy landscape that cannot be accessed in conventional experiments can be probed. However, in order to realize the full utility of the single molecule force spectroscopy it becomes necessary to use force in distinct modes (constant force, force-ramp, and other combinations) along with reliable computations that can mimic the experimental conditions as closely as possible. We conclude the paper with the following additional remarks.

Transition state movements show changes from plastic to brittle behavior: The small size and simple architecture of RNA hairpins has allowed us to explore their response to force over a range of r_f that spans four orders of magnitude. The lowest r_f value is close to those used in LOT experiments. A key prediction of our simulations is that the location of the transition state for P5GA moves dramatically from about 6 nm at low r_f to about 0.5 nm at high r_f (see inset to Fig. 2-E). The large value of Δx_F^{TS} at low r_f suggests that P5GA is plastic while the small Δx_F^{TS} at high r_f is suggestive of brittle behavior. The mechanical properties of RNA structures can be drastically altered by varying the loading rate which is reminiscent of the changes in the visco-elastic behavior of polymeric materials that changes with frequency. The transformation from plastic to brittle behavior can be captured using the fragility index[15] used to describe mechanical unfolding of hairpins. Although we have discussed the r_f-dependent movement of the transition states using P5GA as an example we predict that this result is general and should be observed in other RNA structures as well.

Acknowledgements

This work was supported in part by a grant from the National Science Foundation through grant number NSF CHE-05-14056.

References

1. J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, Jr., and C. Bustamante, *Reversible unfolding of single RNA molecules by mechanical force*, Science, **292**, 733–737, 2001.
2. B. Onoa, S. Dumont, J. Liphardt, S. B. Smith, I. Tinoco, Jr., and C. Bustamante, *Identifying Kinetic Barriers to Mechanical Unfolding of the* T. thermophila *Ribozyme*, Science, **299**, 1892–1895, 2003.
3. C. Bustamante, J. F. Marko, E. D. Siggia, and S. Smith, *Entropic Elasticity of λ-Phase DNA*, Science, **265**, no. 5178, 1599–1600, 1994.
4. J. F. Marko and E. D. Siggia, *Bending and Twisting Elasticity of DNA*, Macromolecules, **27**, 981–988, 1996.
5. K. Kremer and G. S. Grest, *Dynamics of entangled linear polymer melts: A molecular-dynamics simulation*, J. Chem. Phys., **92**, 5057–5086, 1990.
6. C. Hyeon, R. I. Dima, and D. Thirumalai, *Pathways and kinetic barriers in mechanical unfolding and refolding of RNA and proteins*, Structure, **14**, 1633–1645, 2006.
7. C. Hyeon and D. Thirumalai, *Forced-unfolding and force-quench refolding of RNA hairpins*, Biophys. J., **90**, 3410–3427, 2006.
8. D. L. Ermak and J. A. McCammon, *Brownian dynamics with hydrodynamic interactions*, J. Chem. Phys., **69**, 1352–1369, 1978.
9. T. Veitshans, D.K. Klimov, and D. Thirumalai, *Protein folding kinetics: timescales, pathways and energy landscapes in terms of sequence-dependent properties*, Folding Des., **2**, 1–22, 1996.
10. S. Rudisser and I. Tinoco, Jr., *Solution Structure of Cobalt(III)Hexammine Complexed to the GAAA Tetraloop, and Metal-ion Binding to GA Mismatches*, J. Mol. Biol., **295**, 1211–1223, 2000.

11. C. Hyeon and D. Thirumalai, *Mechanical unfolding of RNA hairpins*, Proc. Natl. Acad. Sci., **102**, 6789–6794, 2005.
12. I. V. Hofacker, *Vienna RNA secondary structure server*, Nucl. Acids. Res., **31**, no. 13, 3429–3431, 2003.
13. M. Manosas and F. Ritort, *Thermodynamic and Kinetic Aspects of RNA pulling experiments*, Biophys. J., **88**, 3224–3242, 2005.
14. E. Evans and K. Ritchie, *Dynamic Strength of Molecular Adhesion Bonds*, Biophys. J., **72**, 1541–1555, 1997.
15. M. Manosas, D. Collin, and F. Ritort, *Force-Dependent Fragility in RNA Hairpins*, Phys. Rev. Lett., **96**, 218301, 2006.
16. G. S. Hammond, *A correlation of reaction rates*, J. Am. Chem. Soc., **77**, 334–338, 1953.